



61-199178

Cited Reference No.17 in PCT/IPEA/409

Laid-open Patent Application No. 61-199178 laid open on September 3, 1986

Patent Application No. 60-38648 filed on March 1, 1985

Applicant: Nihon Denshin Denwa Kabushiki Kaisha

Inventor:

Title: Information Input Method

Claim:

1. Information input method provided with:
 - a television camera 1 for picking up the image of a person; and
 - an information input unit 2 for receiving image pickup signal from said camera 1, processing said signal, extracting characteristic quantity from said processed signal and thereby detecting a change in the countenance of said person, in which
 - the detected change in the countenance is transmitted into a calculating machine as command signal indicating the intention of said person.

Other reference numerals in drawing figures:

3...information process unit, 4...upper half part of a person, 5...background, 6...CRT, 7,8...example of menu, 9...cursor, 10...original image, 11...background image, 12...difference image, 13...part of a person, 14...processed image, 15...center of gravity of head, 16...part of face, 17...center of gravity of face

Page(4), left lower column, lines 11-19

The aforementioned method applied to the upper half image of the person can be effectively applied to the image of eyebrows, eyes, a mouth, a nose or ears. The configuration of these components can be extracted and the movement of the components can be recognized sufficiently by using the data that can be obtained with the known image process technique. When the movement of these components is used, the information process unit 3 recognizes the countenance as well as the position information, thereby enhancing man-machine interface.

メウから入力された映像から優先もしくは指示器の位置を抽出するものであるが、あくまで対象はひとコマの静止画像であり、その動きを認識するといふものではないという欠点がある。

また鈴木、岡口、東郷は、「運動検差を利用した3次元表示手法」、昭和59年度電子通信学会総合大会論文発表1541、は本発明と似た主旨の手法の具体例を開示しているが、それは、論文にも述べられているように、コンピュータグラフィックスなどで立体表示を行うとき、表示装置自体がTVカメラなどで観覧者の頭部の3次元位置の動きを常時監視し、その動きに従って表示値を調整させるといふ3次元画像表示手法を提案するものであり、その際、観覧者の動きの抽出には、頭部に付けた3個の発光ダイオードを利用するという方法によつており、頭部画像の利用による方法について、具体的な方法を開示したものでなかつた。

(発明が解決しようとする問題点)
そこで本発明は、人間の動きなどを表情の

⑤ 日本国特許庁 (JP) ⑥ 特許出願公開
⑦ 公開特許公報 (A) 昭61-199178
⑧ 特許請求 未請求 発明の数 1 (全5頁)
⑨ Int. Cl.⁴ 照別記号 庁内整理番号
G 06 K 9/00 C-8320-5B
G 06 F 3/02 Z-7218-5B
3/023 Z-7218-5B
3/03 7165-5B

⑩ 発明の名称 情報入力方式

⑪ 特 願 昭60-38648

⑫ 出 願 昭60(1985)3月1日

⑬ 発 明 者 間 瀬 健 二 横須賀市武1丁目235番地 日本電信電話公社横須賀電気通信研究所内
⑭ 発 明 者 末 永 康 仁 横須賀市武1丁目235番地 日本電信電話公社横須賀電気通信研究所内
⑮ 出 願 人 日本電信電話株式会社 東京都千代田区千代田1丁目1番6号
⑯ 代 理 人 弁理士 並木 昭夫 外1名

て高度な情報入力方式に関するものである。

(従来の技術)

キーボード付ディスプレイ、ライトペン、タブレット、マウス、ジョイスティック等は、従来より広く使用されているコンピュータへの情報入力機器である。

一方、人間の身振り手振りは最も本能的な意志伝達手段として重要な役割をはたしてきている。特に顔の動作は、人間が赤ん坊から幼児へと成長していく過程で、言葉よりも先に使うようになる情報伝達手段である。人間の顔の動作を直接コンピュータで理解できるようにすれば、それは、キーボード等の機器による入力を用い、より自然なマンマシンインタフェースを実現するものとして大きな意味をもつようになると考えられる。また、身体の不自由な人々にも顔面に対する命令伝達手段を提供できることになる。

特許第1144011号(米本、山岸)：「情報入力方式」、特公昭57-1009、特開昭51-10408)などは、手の映像を使い上記問題点を一部解決し、TVカ

を想定し、命令と顔の動作の関係を表1のように定める。動作はいずれも1〜3秒間継続すると考えることができ。

表1

命令	顔の動き
「はい」	頭を上下に一回転する
「いいえ」	頭を左右に一回転する
「ズームアップ」	頭が近づいて来る
「ズームダウン」	頭が遠のいて来る
「左パン」	頭が左を向いて来る
「右パン」	頭が右を向いて来る
「上スクロール」	頭が下を向いて来る
「下スクロール」	頭が上を向いて来る

また、現実の処理を可能とするため、以下の仮定を置く。

- 仮定1「背景は変化しない」。
- 仮定2「顔の基準位置が画像の中心付近にあり、常に画面上の部分が画面内にある」。
- 動きの認識は、(1)顔面画像からの特徴抽出、(2)特徴量による動きの判定、の2段階からなる。

変化として直接映像でとらえ、画像処理して人間の意図(命令)として計算機へ入力可能にするこ

とを、解決すべき問題点としている。

従って本発明は、上述のことを可能にする情報入力方式を提供することを目的としている。

(問題点を解決するための手段および作用)
上記目的を達成するため、本発明による情報入力方式は、人物を対象として得るテレビカメラと、該カメラからの映像信号を入力され処理し、特徴量を抽出することによつて顔記人物の表情の変化を抽出する情報入力装置とから成り、抽出された表情の変化を顔記人物の意図を要する命令信号として計算機へ入力するようにしたことを特徴としている。

以下、具体例を説明すれば次の如く、
画像データを表示しているディスプレイCRTの正面に置いた人物の上半身をCRT上のTVカメラから撮影、表示に対する命令を取り込むことを想定する。

顔面画像の動きから8種類の命令を抽出すること

表1の8種類の命令の抽出を行なうため、(1)顔の大きさ、(2)顔の位置、(3)顔に対する顔の相対位置、を要する特徴量を以下の手順により求める。

(1) 各フレーム毎に顔面画像と背景画像の差分をとり、人物像のシルエットをつくる。

(2) シルエットから頭(首から上)の領域を抽出し、顔の面積と重心を求める。

(3) 顔の領域を抽出し、顔の面積と重心を求める。

上記特徴量に基づき、下記ルールにより動きを分類する。

ルール1：頭が大きく(小さく)なった。

→ 顔が近づいた(遠のいた)。

ルール2：頭が顔に対して相対的に上・下・左・右にある。

→ 顔が上・下・左・右を向いている。

ルール3：顔の向きが連続して上下(左右)に変化した。

→ 「はい」「いいえ」の意味で

マップ、ズームダウンの命令として使う。

次に、頭位置データP1 (x_1, y_1) を頭の重心の位置データと比較して、頭の位置からのずれを調べる。ずれがあれば、そのずれが動いたと判定する。ずれの量が検出されると、頭が動いていると判定する。

さらに、頭の重心位置データP2 (x_2, y_2) を頭の重心の位置データと比較して、頭の回転を調べる。このとき、頭のずれの量を使って修正することによって、頭が移動していても正しい位置を計算できる。頭の回転量は、各映像から得られる頭の重心位置を基準とした頭の重心位置の相対位置の変化をつかうこともできる。頭の回転が、上下あるいは左右に連続して繰り返される

とき、例えば「はい」、「いいえ」と判断する。また、頭の向きを回転量から求めて、これを頭の位置を使うと、大まかであるが視線の方向を知ることが出来る。情報処理装置3は画面の中の注目点を知ることができる。すなわち、例えば、人間が上を向いて注目点が上と認識すれば、情報処理

装置3はそれを上方向カメラ移動と解釈して、第1図のCRT6に表示しているカーソル9を上方向に動かして表示図形7のどこへもつていつたり、画像のさらに上方向を見たいのだと解釈して、画像を下方にスクロールするという処理ができる。またさらに人間が頭を右に向けたのを認識して、カーソルを表示図形8に移動して、そこで、人間が頭に頭を離れば、そのメニューを実行する処理に

はいる、といったようなマンマシンインタフェースを実現できる。

上記の方式を、人物の上半身映像ではなく人物の顔の中にある、眉、目、口、鼻といった点に対して適用しても同等の効果を得ることができる。これらの成分図形の抽出、および動きの認識は、現在の画像処理技術の範囲で得られるデータを用いて十分行なうことができる。これらの成分の動きを利用すると、前記情報処理装置3は位置情報に限らず顔の表情を認識して、マンマシンインタフェースの向上に利用できる。

4. 図面の簡単な説明

第1図は本発明の一実施例を示す説明図、第2図は画像処理の過程を示す説明図、第3図は得られた処理画像の一例を示す説明図、第4図は同じく他の処理画像例を示す説明図である。

符号説明

1...TVカメラ、2...情報入力装置、3...情報処理装置、4...上半身、5...背景、6...CRT、7、8...メニューの例、9...カーソル、10...原画像、11...背景画像、12...差分画像、13...

抽出する。これには、原画像10の背景のうち、処理画像14の値1に対応する部分、すなわち頭の部分だけに對して原画像10を減算し、しきい値処理によつて抽出する。しきい値処理は、例えば人物4の頭が黒く、頭が白っぽいという性質を使って、あるしきい値より明るい部分を顔(値2)として判定する。第4図の16が顔として判定された部分である。そこで、顔の部分16の重心17のデータ($P2 = x_2, y_2$)を計算する。

情報入力装置2は上記の計算をTVカメラがとらえた映像ごとに行い、得られる頭の位置、頭の重心位置座標、頭の重心位置座標を逐次出力する。情報処理装置3は、情報入力装置2から逐次受け取る前記のデータと、最初に受けとつて位置データとを処理する。まず頭の位置データと常時調べて増加しているか、減少しているかをチェックし、それぞれのとき、頭が近づいている、あるいは遠のいていると判定する。画像を表示しているときは、これらの判定結果をたとえばズー

ムを表現できる。

装置3はそれを上方向カメラ移動と解釈して、第1図のCRT6に表示しているカーソル9を上方向に動かして表示図形7のどこへもつていつたり、画像のさらに上方向を見たいのだと解釈して、画像を下方にスクロールするという処理ができる。またさらに人間が頭を右に向けたのを認識して、カーソルを表示図形8に移動して、そこで、人間が頭に頭を離れば、そのメニューを実行する処理に

はいる、といったようなマンマシンインタフェースを実現できる。

上記の方式を、人物の上半身映像ではなく人物の顔の中にある、眉、目、口、鼻といった点に対して適用しても同等の効果を得ることができる。これらの成分図形の抽出、および動きの認識は、現在の画像処理技術の範囲で得られるデータを用いて十分行なうことができる。これらの成分の動きを利用すると、前記情報処理装置3は位置情報に限らず顔の表情を認識して、マンマシンインタフェースの向上に利用できる。

4. 図面の簡単な説明

第1図は本発明の一実施例を示す説明図、第2図は画像処理の過程を示す説明図、第3図は得られた処理画像の一例を示す説明図、第4図は同じく他の処理画像例を示す説明図である。

符号説明

1...TVカメラ、2...情報入力装置、3...情報処理装置、4...上半身、5...背景、6...CRT、7、8...メニューの例、9...カーソル、10...原画像、11...背景画像、12...差分画像、13...

特開昭61-199178(3)

像を連続入力し、該動画像を情報入力装置2において画像処理し、頭位置データP1 (x_1, y_1)、および頭の重心位置データP2 (x_2, y_2)を、x、y座標に基いて計算して、結果をカメラと情報処理装置3に出力する。

情報処理装置3は、上記のデータ(座標や重心位置)の変化などから、まず人物4の頭がTVカメラ1に近づいているかあるいは遠のいているか、さらに頭が上下左右のどの方向に動いているか、また頭(顔)が上下左右のどの方向をむいているかを検出する。さらにこれらの動きの速さ(単位時間あたりの向きの変化の回数など)から、頭が振られているかどうかなどを検出する。その後、以上の信号から前記の頭の動きによつて変化する検出された命令に変換する。このとき、動きに対する意味付けは任意の方法があり度換手続は限定されない。

情報入力装置2にはあらかじめ背景の画像を取り込み該装置内に格納しておく。また、人物の上半身4の位置を決定するため、人物がTVカメラ

差分画像12)は、背景の画像は0に近く、その他の画像は比較的大きな値となる。但し、上半身画像部分は、そこから対応した位置にある背景画像分を差し引かれるので、一律的な画像にはならないことがある。

次いで背景画像を、あらかじめ与えておいたしきい値で2値化すると、(エ)に示す如く人物の部分13が抽出できる(しきい値以上を値1、それ以下を値0とする)。頭の部分だけを処理対象とするために、頭から下の部分を削除する。これには、首のところで値1の図形(人物部分)の幅が小さくなることをつかって首の位置を検出する。この時点での処理画像14は第3図に示すうになつてゐる。

次に、情報入力装置2はまずこの処理画像14から値1の図形の面積を計算する。これが頭の面積データSとなる。さらに、値1の図形の重心位置を既知の手法により計算する。これが頭の重心15のデータ($P1 = x_1, y_1$)となる。次に処理画像14と原画像10を使って頭の部分を

ある。
このようにして本発明は、例えば頭の動きを調べし、その画像を簡単な方法で処理して認識し、計算結果を入力することを可能にしている。

(実施例)

次に図を参照して本発明の実施例を説明する。第1図は本発明の一実施例を示す説明図である。同図において、1は人物の上半身映像を取り込むTVカメラ、2は取り込んだ映像を処理して頭の位置情報、動作情報を出力する情報入力装置、3は動作情報から前記の命令などを判定して、該命令を利用する情報処理装置である。また図中、4から9は動作を説明するためのものであり、4はTVカメラで捉える人物の上半身、5は背景、6は該人物が例えば本発明による方式を使って画面上的メニューを選択するとしたときにメニューを表示するCRT、7、8はそれぞれメニューの例、9は画面上的カーソルである。

次に動作を説明する。まず第1図に示すような配置におかれたTVカメラ1から上半身4の動

1の正面に座つたときに、情報入力装置2および情報処理装置3に信号を送り、その時の前記データすなわち頭の座標、頭の重心位置P1、頭の重心位置P2を位置データとして記憶しておく。本方式では、これらの背景と位置データを使

つて、頭の動きを判定する。これらの背景と位置データに関するデータは基本的な使用では固定式的だが、場合によつて、背景が変化したり、頭の位置が変わることもあるものでこのときは、随時、背景の画像と位置データの平均値を求めて平均値を使うとか、大きな差が生じた場合には更新して新しいデータを使用する。

情報入力装置2は次のように動作して、頭の位置データSと、重心データP1、P2を計算し出力する。第2図を用いて説明する。

まず、TVカメラ1から入力した(イ)に示す上半身像(原画像10)から、前記記憶しておいた背景画像11(ロ)に示す、差し引いて得た差分画像12(ハ)を計算する。この差分は、画像の各画素ごとにそれぞれ行う。ここで得られた(ハ)に示す画像(

